From Greetings to Corruption: Politicians, Political Parties, and Tweeting in India

Lia Bozarth
University of Michigan
Ann Arbor, United States
lbozarth@umich.edu

Ceren Budak
University of Michigan
Ann Arbor, United States
cbubak@umich.edu

Anmol Panda
Microsoft Research
Bangalore, India
anmolpanda07@gmail.com

Joyojeet Pal
Microsoft Research
Bangalore, India
joyojeet@gmail.com

ABSTRACT

We present an in-depth, language-centric, large-scale study of topical discussions on Twitter by 1,711 Indian politicians from 20 competing political parties in a 4-year timespan. We first show that politicians of all parties collectively indulge more in establishing personal branding through low-substance, personality-focused messaging as opposed to broadcasting policy stances. Additionally, compared to the party-in-power, opposition politicians collectively post more complex tweets and demonstrate higher negativity (e.g., using person-based attack hashtags) especially regarding the issue of corruption. Finally, through a contextual examination of the most retweeted messages from two key leaders—the prime minister and the leader of the largest opposition party, we find that there are qualitatively important distinctions between their styles—while the former focuses on positive-themed messaging, the latter employs confrontation and aggressive language with direct attacks on individuals and issues.

CCS CONCEPTS

• **Human-centered computing** → **Social media**; *Social content sharing*; *Computer supported cooperative work.*

KEYWORDS

online political communication, social media and political communication, the Global South, Indian politics, Indian politicians on Twitter

1 INTRODUCTION

Information & Communication Technology (ICT) has fundamentally altered the social and economic lives of people in the Global South. Its significance can be seen in business transactions, healthcare, personal safety, education, to name a few. In the past decade, it has also become increasingly clear that ICTs including social media have impacted the content and channels of political communications in ways that have dramatically changed both campaign and non-campaign outreach [5, 42, 51]. New forms of rhetorical construction, which allow politicians to control what subjects they address, facilitate both the personalization of communication at the cost of policy-relevant discussions, and create a reward mechanism through tech-enabled rapid message dissemination [28, 52, 53]. Thus far, research has been largely centered on US and European adopters of social media. These studies have examined Western politicians' social media strategies on issues of brand building, and championing for or against salient legislative issues [16, 34]. In recent years, the growth of access to low-cost mobile devices has brought social media to center-stage for political communication in various parts of the world—with major world leaders such as Joko Widodo, Jair Bolsonaro, Rodrigo Duterte, and Narendra Modi all gaining reputations as masters of campaigning using Twitter.

Compared to Western societies, India bears a significantly distinct news media and political landscape. For instance, work [65] has shown that political journalists still treated public officials with a colonial legacy of deference, though the trend had started to reverse with the increase of vernacular journalism [65]. Additionally, there are over 20 competing political parties in India and, unlike the United States, party-ideology does not neatly fall into the liberal versus conservative spectrum. Rather, it is interwoven with religion, region, language, caste, and race complicated by India's long history. Lastly, in India, candidate selection and candidates' presentations of themselves on social media are typically managed and guided by parties and parties leaders [64].

Motivated by these unique circumstances, we set out to explore the characteristics of Indian politicians' online communication. Our work is primarily focused on examining the Bharatiya Janata Party (BJP), and the Indian National Congress party (INC). The former is the current governing party of India, it's a right-wing party with policies that reflect Hindu nationalist positions [18]. The later is

ideologically center-left and traditionally follows democratic socialist philosophy [18]. In our work, first, we use semi-supervised methods to populate a list of 1711 Indian politicians on Twitter and acquire 4.6 million tweets from their accounts. We then create two bins of topics: i) policy-relevant topics which typically appear on election manifestos, and ii) non-issue topics. For our paper, we select `corruption`, `development`, `inflation`, `technology`, `poverty` as policy-relevant topics [1] and `greetings` for non-issue topic [2]. We annotate each applicable tweet into one of the aforementioned 6 categories of tweets. Finally, we examine the distribution of topics political elites choose to discuss online (e.g., what's the most frequently discussed issue?), and study the linguistic differences of such discussions (e.g., which topic is imbued with the most polarizing sentiment?). Our paper makes the following contributions:

- To the best of our knowledge, this is the first large scale study that uses both manual and automated approaches to identify thousands of politicians in an electoral system in a Global South setting. As such, this valuable dataset will be shared with the research community in the future to further advance non-Western political communications studies.
- We show that politicians more frequently tweet casual outreach subjects compared to political manifesto-relevant topics. This suggests that elites see value in signaling relatability through casual talk rather than lean more heavily towards a policy-centric message. Further, BJP clearly dominates other parties in both social media presence and tweets of soft topic `greetings`.
- We also observe that the ruling BJP party is relatively less engaged in antagonistic messaging — it has more engagement with sports-related and nationalistic hashtags. Further, tweets by BJP politicians are also more simplistic on average. In comparison, the opposition INC messages relatively more frequently on issue-centric and antagonistic hashtags aimed at the governing party. Non-governing parties were also found overall to have more complex tweets as well as more negative sentiments in their tweets.
- Finally, by qualitative examining tweets of key leaders, we demonstrate that tweets about "policy" are sometimes confrontations or political performance rather than substantive discussions. Further, opposition leaders are often rewarded with for being aggressive on social media.

These findings help further our understanding of both the affective logics that politicians aim to employ in reaching out to citizens, and offer a view into the ways political parties in the largest democracy in the world organized its media outreach efforts around specific narratives. Using text-based analysis, we demonstrate that not only are political elites collaborating with each other to build a common narrative for specific policy issues, they are also collectively building a "personality"—using party leaders as exemplars—for their own parties.

---

[1] Although there are other "policy-relevant" topics, such as defense, security, federalism, citizenship, etc. we stuck to issues that are not confounded by nationalistic or polarizing rhetoric, which can end up inflating metrics of resonance such as retweets or likes on Twitter, which we examine for effect.

[2] Non-issue tweets include those related to birthdays or festival greetings, condolences, etc. While messaging on such 'non-issue' tweets serves an important purpose such as contributing to personal branding, these are less likely to be immediately policy- or election manifesto-relevant.

2 RELATED WORK

The enthusiastic embrace of social media that followed spontaneous and affect-driven events such as the 'Arab Spring' [46, 54, 74] has since been tempered by the possibility that social media actually enhanced established political actors' ability to systematically use social media to monitor citizens [22, 33, 67, 78], or to feed them a preferred narrative in an organized and sustained way [21, 56]. In this paper, we seek to understand politicians' and their parties' choices of what to talk about, and, to a lesser extent, what not to, in a medium that is primarily unidirectionally broadcast - in that politicians themselves do not engage with follow-up content such as comments, retweets etc, but may benefit from them in terms of message propagation. Here, we first discuss the changing media environment and its impact on political journalism [5, 11, 30, 51, 64, 66]. We then survey literature on individual politician's communication strategies [2, 4, 20, 45, 45, 47, 61]. Finally, we review studies that address how political parties advance party-based agendas [17, 28, 34].

New Media Environment: Individual politicians' strategies are particularly tied to the prevalent practices of journalism, and how these may enable certain personalization practices. With changes in the mediatization environment, independent journalists are less enabled to force politicians to address topics of traditional mainstream media interest [48] while partisan journalists are better enabled to promote a politician's narrative both online and offline [11]. Structural changes in the media ecology brought about by online readership intersect with the ability of social media to enable the amplification of biases, making it profitable for journalists to toe a majoritarian line [75]. Focusing on India, there have been growing concerns regarding the state of media freedom since the start of the Modi-led BJP government in 2014 [64]. Commentators have further noted the intersection of social media with the emergence of sensationalism in the news in India, where investigative or probing journalism has been pushed towards theatrical events such as sting operations [73]. Issues of public importance such as corruption and development are mediated through the performance of outrage by Indian celebrities at these events on social media instead [66].

These studies suggest that the growth of social media as a source of news has exacerbated the challenges with quality news access, creating a heteronomous effect on journalistic practices, essentially enabling political hegemons to dominate the media discourse [49]. Given these considerations, researchers including Bennett and Pfetsch [7] stressed the importance of examining and understanding the changes in political discourse within the new media environment, and how these changes affect democratic institutions at large. Luckily, the digital prints, such as policy tweets on Twitter, left by these political entities [77], provide interested parties opportunities to study individual politicians at an unprecedented scale and depth.

Social Media and Personalized Communication by Political Elites: The subjects a politician chooses to address on their public Twitter feed directly relates to the brand image they choose to project. While much of the early work focused on Western societies, there is a growing body of work that examines the role of social media in the communications of individual politicians in the Global South [2, 4, 20, 45, 45, 47, 61]. Within the context of India, Jaffrelot [37], for instance, notes that "the second-largest independent PR firm in America, charged over $25,000 a month to manage Modi's

[media] account", underscoring the importance of digital branding to Prime Minister Narendra Modi. Additionally, Pal et al. [59] observe that politicians selectively put out policy-relevant or personal brand building oriented content in tweeting depending on the stage of the electoral cycle. Furthermore, political elites are not only using data mining techniques to classify potential supporters into granular groups and then personalize their campaign messages targeted at these groups accordingly, but these messages are also becoming increasingly focused on "lifestyle" messages instead of being policy and issue-based [6, 8, 35, 43]. Finally, in addition to personal branding, recent work [60] also demonstrates populists' ability to leverage social media to attack and discredit their opponents and erode the goodwill and opportunities of partnership between both citizens and political elites from competing parties.

Overall, studies here highlight the movement of online political discourse towards being more personal, affective yet trivial, rather than civil and informative.

Party-based Collaborative Strategies on Social Media: On one hand, a large body of work examining the ways social media has led to the personalization of politics, leading to individual personalities playing a more important role in party politics [14, 17, 28, 62]. However, parties offer structural support, professional brand management, and the potential of collaborative online action through their online networks. One of the early systematic studies of two-party online strategy by Hemphill et al [34] showed that Republican and Democratic officials collectively used distinct hashtags on Twitter to frame or counter-frame policy-relevant discussions.

Active collaboration using hashtags can reinforce individual politicians' stances or garner momentum for discussions on these issues. Studies also show that members of a party can perform collective antagonistic action and that this depended on their position of power. As an example, studies of the 2008 and 2012 US presidential elections have shown that Republicans in both cases— McCain and Romney—focused on attack, while Democrat Obama focused on acclaim, with parties in each scenario fell behind the narrative of the leader [10]. The 2012 campaign of Mitt Romney has also been used as a case study to propose that a second-term incumbent is likely to be an object of attack from an opposition party [29]. Additionally, political elites are more likely to interact with others from the same party [4, 26] than those from opposing parties, in an attempt to increase each others' visibility [26]. Moreover, Boggild and Pedersen [9] contend that while many recent studies indicate that politicians are using a more individualized approach in campaigns and elections, politicians from systems in which parties have direct control over the nomination process are less likely to engage in personalized campaigning. Thus, in India where candidate selection is typically managed by parties, social media usages by members of the same party are potentially more uniform due to centralized party-level outreach management, and the interactions or collaborations between same-party politicians more significant.

The studies listed above are all valuable contributions that provide deeper insights into how social media have fundamentally altered the media environment and political discourse online. Unlike these studies which are predominately Western-centric or focused only on a small set of key politicians, here, we dive deep into the

what and how of Indian politicians' speech on Twitter [3]. We leveraged Twitter networks to classify a substantially large number of influential Indian politicians both at national and regional levels. Our study has a dataset spanning 4 years and we use both qualitative and quantitative methods to examine social media strategies of these political entities.

3 DATA

Our dataset consisted of 4.6 million tweets [4] in both English (2.94M) and Hindi (1.66M) from Jan, 2014 to October, 2018, contributed by 1711 Indian politicians from both state and national levels. We note that 779 accounts were Bharatiya Janata Party (BJP) politicians who supplied 2.34M tweets (or, 51.0% of total tweets); 417 accounts were from the principle national opposition the Indian National Congress (INC) politicians with an aggregated sum of 1.25M tweets (or, 26.9% of total tweets); and the remaining 515 politicians from other parties posted 1.02M tweets (or, 22.1% of total tweets). The data is thus significantly skewed towards BJP politicians. This underlines the party's dominance of social media, partly led by its leadership's diktat that only politicians with over a certain threshold of followers on social media could be offered party seats for elections [12].

The aforementioned 1.7K politician Twitter accounts are identified using a mixed-methods approach. Furthermore, we also used a cosine similarity-based unsupervised process to assign tweets into the following 7 non-exclusive categories: 1) information and communication technologies (`technology`) related tweets, 2) poverty & welfare (`poverty`), 3) economics and development (`development`), 4) corruption, scams, and bribery (`corruption`), 5) inflation and fuel price surge (`inflation`), 6) greetings and holiday celebrations (`greetings`), and 7) other (`other`). Both of these classification procedures are described below.

3.0.1 Twitter Account Classification. We used both manual and supervised learning methods to generate a comprehensive list of Indian politicians on Twitter. Our classification covered a considerable number of politicians from 20 distinct parties (both national and regional) including BJP, INC, Shiv Sena, CPI(M), AAP, etc.

Manual Approach: We used Twitter's native search function to query for politician Twitter accounts using the names of elected parliamentary representatives, key party officials (such as party presidents, general secretaries), chief ministers of states, and state legislators. We manually looked through the top 20 results returned for each query and identified all accounts matched to actual politicians. This procedure gave us a total of 1002 manually collected political handles.

Supervised Learning Method: We used a 2-step process to identify accounts of additional Indian politicians. First, we built a well-labeled dataset of 1002 known politicians (positive class) and the

---

[3]While there are several popular social media platforms all of which garnered significant interest from the academia, in the paper, we focused on Twitter due to the following reasons. Twitter has had early success in political use by key politicians, and has been used widely by various political parties since the 2014 general elections [3]. Additionally, Twitter, as pointed out by prior work [44], differs from other social media platforms due to its short effective diameter and non-power-law follower distribution. In other words, Twitter is ideal for information broadcasting. Consequently, much research has been done to examine the adoption of Twitter by political elites. See [39] for a review. Yet, in-depth studies that focus on the Global South are still lacking.
[4]We use *Tweepy*, a python library, to continuously stream tweets and store the results in a remote ElasticSearch instance.

top 800 retweeted non-political accounts (negative class). We then trained a `LogisticRegression` binary classifier adopting the common Bag-of-Words Ngram-based approach [38] using user profile description text [5]. We achieved a precision score of 0.967, a recall score of 0.659 and f1 score of 0.762. In the second step, we aggregated all Twitter accounts that are followed by at least one politician in the labeled data. We then used our classifier to label these additional accounts. For accounts that are classified as politicians, we keep those that meet at least 1 of the following criteria: i) has contributed more than 1000 tweets, ii) has more than 1000 followers. Finally, we manually examined all remaining accounts and discarded all false positives.

3.0.2 Tweet Classification. Keywords and/or Cosine similarity measurement based document clustering has been used extensively by many prior studies [1, 19, 72] focused on political communication on social media. In this paper, we used a similar approach to cluster tweets into different issue-based and personal-appeal-based categories (details in the Appendix).

Briefly, we generate a numeric vector for each unique word. For instance, the word "govt" can have the corresponding vector of [-1.9078784 1.654422 1.8524699...]. Vectors of words that share more similar contexts are closer together in space (i.e. having higher cosine similarity score). Here, we first manually selected a handful of keywords related to `technology`, such as "technology", "digital", et cetera, and then used cosine similarity scoring to discover additional keywords that share comparable semantic meanings to the selected words. Next, for each tweet, we categorized it as `technology` if it contains 1 or more `technology` keywords. We repeated the process for other categories. Using this approach, we generated 157 keywords for `technology`, 91 for `poverty`, 131 for `development`, 151 for `corruption`, 37 for `inflation`, and 275 for `greetings`. Due to the overlapping nature of political discourse on Twitter, some tweets were classified under multiple themes. Though, if a tweet has no matching keywords, it's assigned to `Other`. Table 6 in the Appendix contains the complete lists of keywords.

Using this approach, we labeled 187.7K or 3.4% tweets as `technology`, 188.1K or 3.4%, as `poverty`, 568.1K or 10.2%, as `greetings`, 327.3K or 5.9% as `development`, 171.0K or 3.1% as `corruption`, 63.2K or 1.1% as `inflation`, and the remaining `Other` tweets. This distribution of tweets suggested that politicians generally tweet more about holiday celebrations and greetings in comparison to actual policies related to technology and development. In order to assess the performance of our classification, for each category of tweets, we randomly select a representative sample (e.g. we calculate sample size using 95% confidence level and ±3% confidence interval) and manually assess whether each tweet indeed focuses on technology, poverty, and/or development related topics. We see an 80.3% precision for `technology`, 82.0% for `poverty`, 90.0% for `development`, and comparable results for others. Finally, we aggregate all categories of tweets together and observe a combined precision and recall of 85.4%, and 82.0% respectively.

4 ANALYSES

A look at the overall descriptive statistics from the sample gives a sense of the landscape of political twitter in India. Figure 2(a) shows that Prime Minister Narendra Modi was clearly the dominant figure both in terms of his following, and the extent to which his tweets got retweeted. However, there are several other political leaders with significant presence and influence online. For instance, the main leader of the rival INC party, Rahul Gandhi, gets more average retweets than Modi despite the later currently has several times as many followers. We also see a trend that all of the accounts with above-500-median-retweets were leaders of specific parties (e.g. Akhilesh Yadav of the Samajwadi Party, Arvind Kejriwal of the Aam Aadmi Party), even if they weren't necessarily the most followed leaders from within the overall sample. This suggests a centralizing tendency for social media communications to coalesce around party leaders. Additionally, given that recent work observed high bot activities on Twitter [76], we also randomly sample a subset of (sample size corresponds to 99% confidence-level and +/-3 confidence interval) Narendra Modi and Rahul Gandhi followers and then use Botometer [24] to determine their bot scores (technical details are in the Appendix). We observe that 63% and 54.7% of randomly sampled Modi and Gandhi followers are bots. These numbers are much higher than the reported 9% to 15% by another study [76], suggesting high follower number inflation which occurs across parties [6].

The Analyses section is structured as follows. We first assess the distribution of original tweets (i.e. not including retweets) for each category of tweets (e.g. `greetings`, `inflation`, etc.) across different parties. We then examine how politicians differ when talking about policy-related and non-substantive issues through the lens of i) hashtag usage, ii) text complexity, and iii) sentiment. Finally, we focus on qualitatively assess Twitter strategy differences of key party leaders: Narendra Modi and Rahul Gandhi.

4.1 Frequency Distribution of Direct Contributions for Each Category of Tweets

Results are summarized in Figure 2(b). We observe that BJP contributed 348.9K `greetings` tweets, accounting for 12.8% of its total tweets. This is much higher than the INC which only contributed 124.3K or 8.1% of its total tweets to `greetings` (similar observation for politicians in 3rd parties). These numbers suggest that the governing party members are more frequently engaged in non-issue-based greetings (over 50% more than the main opposition party). Moreover, BJP also demonstrates a higher preference to discuss `development` related issues such as economic development and GDP. On the other hand, INC shows a higher affinity in discussing `corruption` and `inflation`: 83.5K or 5.4% of its total tweets are focused on `corruption` while the number is only 2.2% for both BJP and 3rd-party politicians. Similarly, the median number of tweets by BJP accounts on `technology` is 83 and `development` 167, significantly higher than those of INC (48 and 107). In comparison, the median number of tweets posted by INC on `poverty` is 101, `inflation` 46, and `corruption` 118, whereas the numbers are 76,

---

[5] For instance, words including "mp", "minist", and "parliamentari" in user profile have the largest positive weights, while the words "super", "andhra", and "columnist" have the largest negative weights.

[6] Another possible explanation is that Botometer has a higher false-positive rate for professionally managed organizational accounts [24].

14 and 48 for BJP. Results suggest that opposition parties use social media more extensively for critical messaging aimed at the governing party.

Figure 1: Overview of Politicians' Social Media Presence.
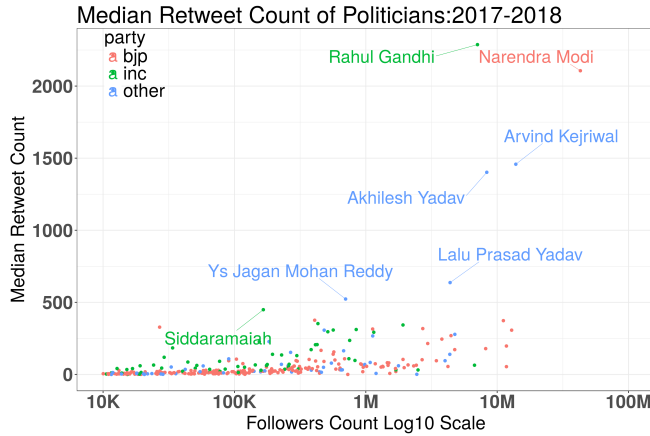
(a) Median Retweet Count for Individual Political Account
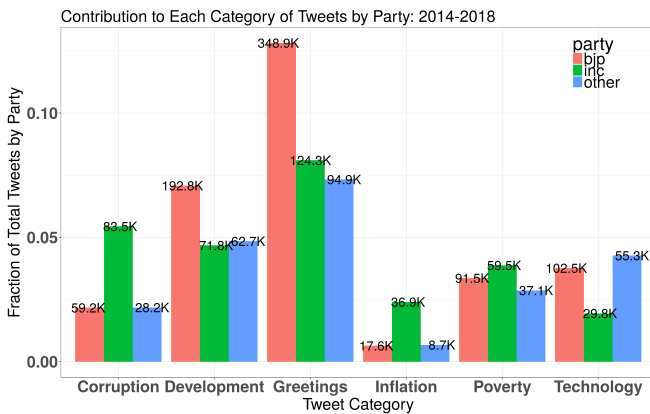
(b) Party-level Contribution by Tweet Categories

The top 10 contributors for each category of tweets determined by the absolute number of tweets are shown in Table 1. Unsurprisingly, politicians associated with the tech industry such as the ministers of Electronics (RS Prasad), Science and Technology (Harsh Vardhan), and Commerce (Suresh Prabhu) contributed significantly to the Twitter conversation on `technology`, as well as Andhra Pradesh Chief Minister Chandrababu Nadu, and INC member, tech billionaire Nandan Nilekani. On `development` discussions we intersections with the most vocal leaders on issues of technology and poverty. Dharmendra Pradhan, Piyush Goyal and Suresh Prabhu who featured in the `technology` list also contributed significantly in discussing `poverty`, suggesting that these are co-occurring themes, an idea that has purchase in a body of *ICTD* work that proposes technology as fundamentally tied to the idea of development in the Global South.

In the `greeting` related tweets, Prime Minister Narendra Modi emerges as the leader, and all the highest tweeting politicians in this set are from the ruling party. This aligns well with previous research which suggests that Narendra Modi's online presence aims to build an image of national harmony [40], but also shows that parties in power are more likely to highlight a positive discourse. The top contributors of `corruption` are all key party spokespersons, rather than the top-ranking leaders themselves, suggesting that the less savory topics are left to be addressed by party mouthpieces. Also, as expected, opposition figures are more likely to engage with it – the main campaign of the INC against the BJP was eventually on the issue of corruption. The discussion on `poverty` is fairly diverse in terms of the actors involved—a significant presence of key opposition figures (Arun Yadav, Sitaram Yerchuri, Ashok Gehlot), and high contributions by the central government ministers who deal with issues of poverty and rural development as part of their office requirements.

The discussion on `inflation` is also dominated by non-BJP politicians. However, the list of biggest contributors to the subject is led by Dharmendra Pradhan, who is a false positive due to his position as petroleum minister (he dominates the list by repeatedly tweeting about petroleum which is otherwise typically an inflation-related subject), and Sanju Verma, a BJP economist whose tweets are largely about defending the party against inflation-related charges. The lukewarm tweeting of BJP members on inflation and corruption reaffirms that the party in power is less inclined to engage with negative topics.

Overall, results here are comparable to work on the topical spread of politicians' tweets by Hemphill et al. [34], who find that about 9% of US Congresspeople's tweets are on soft topics (2% thanks or regards and 7% on telling a story about their day), while the two major parties in India - the BJP scores in greetings at 12.8% while the opposition INC has about 8.1%.

Table 1: Top 10 Contributors (Followers>=100K) by Absolute Tweet Count for Each Category of Tweets

Category	Top10.Contributors
Technology	DIPP India(other 2073); Ravi Shankar Prasad(bjp 1977); Amitabh Kant(other 1696); Dr. Harsh Vardhan(bjp 1550); Suresh Prabhu(bjp 1252); Dharmendra Pradhan(bjp 1075); Jayant Sinha(bjp 908); Narendra Modi(bjp 896); Manoj Sinha(bjp 861); data.gov.in(other 796)
Poverty	Agriculture INDIA(other 1991); Yogendra Yadav(other 1297); Radha Mohan Singh(bjp 1182); Jitu Patwari(inc 877); ShivrajSingh Chouhan(bjp 828); Raghubar Das(bjp 785); Narendra Modi(bjp 573); UP Congress(inc 565); Narendra Singh Tomar(bjp 557); BJP Delhi(bjp 536)
Greetings	Narendra Modi(bjp 2762); Kailash Vijayvargiya(bjp 2647); Sudarsan Pattnaik(other 2125); Raghubar Das(bjp 2091); Mukhtar Abbas Naqvi(bjp 2012); Sidharth Nath Singh(bjp 1784); Nand Kishore Yadav(bjp 1776); Ashwini Kr. Choubey(bjp 1741); Vasundhara Raje(bjp 1687); Yogi Adityanath(bjp 1617)
Development	Raghubar Das(bjp 2180); Suresh Prabhu(bjp 1922); Agriculture INDIA(other 1852); Dharmendra Pradhan(bjp 1485); Radha Mohan Singh(bjp 1308); Narendra Modi(bjp 1259); Vasundhara Raje(bjp 1229); CMO Chhattisgarh(bjp 1228); Piyush Goyal(bjp 1118); Dr. Pankaj Shukla(bjp 1078)
Corruption	Sanjay Jha(inc 1277); Tejashwi Yadav(other 833); Subramanian Swamy(bjp 765); BJP Delhi(bjp 679); Jitu Patwari(inc 655); Sanjay Singh AAP(other 644); Sanjay Nirupam(inc 638); Sambit Patra(bjp 607); Amit Malviya(bjp 576); Randeep Singh Surjewala(inc 552)
Inflation	Dharmendra Pradhan(bjp 839); Jitu Patwari(inc 307); Sanju Verma(bjp 277); Sanjay Nirupam(inc 244); Alka Lamba(other 238); Randeep Singh Surjewala(inc 229); Harish Rawat(inc 215); Sanjay Jha(inc 202); Sitaram Yechury(other 159); Yogendra Yadav(other 154)

4.2 Linguistic Attributes Comparison Across Parties

Here, i) we first analyze the most commonly used hashtags by each political party given that prior research observes that parties strategically use hashtags to frame or counter-frame specific narratives on controversial issues [34]. ii) We then determine the difference in language complexity between tweets of substantive topics and `greetings` tweets. Finally, sentimental appeal in political communication has important consequences (e.g. altering voter behavior) [68, 71]. Thus, iii) we also conduct sentiment analysis on each category of tweets to evaluate politicians' strategic use of emotional appeals and assess whether party-based differences exist.

4.2.1 Hashtag Usage. In this section, we focus on BJP and INC politicians' use of hashtags in `greetings` and `corruption` tweets: the former category is a non-issue while the later is a very contentious topic in India [13].

First, we hand-coded each of the top 50 hashtags from either BJP or INC under the thematic heading of `greetings` into 5 subtypes: i) party-based greetings (such as birthday greetings to party members), ii) nationalistic greetings (e.g. greetings for soldiers on independence day), iii) sports-related, iv) Hinduism-related, and v) generic. We then apply the following regression model to show the usage likelihood of these hashtags as a comparison between BJP and INC. First, we write H as the entire set of hand-coded hashtags, and $P_{greetings}$ as accounts who posted at least 1 `greetings` tweet. For a given hashtag $h \in H$ and politician $p \in P_{greetings}$, we assign dependent variable $y_{p,h} = 1$ if p used h at least 1 time, and $y_{p,h} = 0$ otherwise. We denote $type_h$ as h's subtype, and $party_p$ as p's party. Further, we write $status\_count_p$, $follower\_count_p$, and $friend\_count_p$ to control for p's overall Twitter activity level. Finally, we apply logistic regression equation:

$$y_{h,p} = \beta_0 + \beta_1 * type_h * party_p + \beta_2 * status\_count_p + \\ \beta_3 * follower\_count_p + \beta_4 * friend\_count_p \quad (1)$$

As shown in Table 2, INC politicians are significantly less likely to tweet with `greeting` hashtags as a whole. Additionally, the likelihood of a nationalistic hashtag being used by a politician is reduced by a log odds of 1.36 if the politician is from INC. We can therefore conclude that the use of nationalistic greetings as well as shout-outs to sports are a consistent part of the brand image of the BJP in the period studied. Further, we also see that INC party members are less likely to greet each other on social media compared to BJP. This is consistent with our prior observations: the BJP party as a whole, lead by Narendra Modi, is more invested in branding itself as, warm, optimistic and personable.

Similarly, we manually labeled each of the top 50 hashtags from either BJP or INC in `corruption` into 3 subtypes: i) personal (targeting an individual or a group), ii) issue-based (targeting a particular incident or issue), and iii) generic (neither personal nor issue-specific). We observe that BJP is significantly more likely to talk about corruption using generic terms #SaafNiyatSahiVikas (Clean Intent Good Development), #AntiBlackMoneyDay, #IndiaFightsCorruption, and #BlackMoneyCrackdown. These hashtags address corruption but do not have an explicit target or villain. In

Table 2: BJP and INC Politicians' Likelihood of Using Various Subtypes of Greeting-related Hashtags

	Dependent variable:
	had_used_hashtag
Hashtag Type	
hindu	−0.172* (0.093)
national	−0.405*** (0.110)
party	−0.623*** (0.107)
sports	0.073 (0.100)
Party	
INC	−0.383*** (0.114)
Hashtag Type:Party	
hindu:INC	−0.077 (0.181)
national:INC	−1.357*** (0.335)
party:INC	−0.700*** (0.254)
sports:INC	−0.300 (0.208)
Followers Count	0.107*** (0.037)
Friends Count	0.092* (0.051)
Statuses Count	0.283*** (0.052)
Constant	−6.373*** (0.240)
Observations	168,597
Log Likelihood	−6,680.171
Akaike Inf. Crit.	13,386.340
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 3: BJP and INC Politicians' Likelihood of Using Various Subtypes of Corruption-related Hashtags

	Dependent variable:
	had_used_hashtag
Hashtag Type	
issue	−1.700*** (0.127)
personal	−2.471*** (0.147)
Party	
INC	−0.238*** (0.091)
Hashtag Type:Party	
issue:INC	2.199*** (0.159)
personal:INC	1.488*** (0.193)
Followers Count	0.189*** (0.040)
Friends Count	0.187*** (0.053)
Statuses Count	0.388*** (0.055)
Constant	−7.080*** (0.255)
Observations	161,700
Log Likelihood	−6,104.820
Akaike Inf. Crit.	12,227.640
Note:	*p<0.1; **p<0.05; ***p<0.01

contrast, the INC is significantly more likely to tweet about corruption using hashtags that are issue-centric such as #GreatRafaleCoverUp or #ModiRafaleLiesExposed, both of which refer to the Rafale defence deal which has been a center of a corruption controversy involving the prime minister's office. INC politicians also use hashtags that are personality-centric such as #ModiRobsIndia, or #ChowkidarChorHai (The Watchman is the thief), both of which

are direct personal attacks on the prime minister, Narendra Modi, implying that he himself is corrupt. Indeed, as shown in Table 3 (regression results generated using a comparable model), INC politicians are significantly more like to use both `issue` and `personal` corruption hashtags.

To summarize, our results on hashtags extend prior work focused on the West: similar to the U.S. political actors, Indian politicians are also strategically using hashtags on Twitter for brand and narrative building purposes [34, 58]. Here, the ruling BJP party is using optimistic and nationalistic hashtags to market itself as a party that can move India forward. In comparison, INC is more focused on advancing the narrative that BJP is corrupt and unfit to govern. In other words, our results suggest that not only are established political elites with comparable agendas leveraging social media to collectively broadcast a narrative or counter-narrative for policy-relevant discussions, they are also collaboratively building a "personality" for their party.

Table 4: Regression Results for Readability and Sentiment with Respect to Party and Tweet Categories. Readability is Measured by Flesch, Gunning, and Smog Index. Sentiment is Measured Using VADER. *Insignificant results are not included.*

	Dependent variable:				
	Flesch	Smog	Gunning	Pos Emo	Neg Emo
	(1)	(2)	(3)	(4)	(5)
Party					
INC	0.620***	0.371***	0.367***	-0.030***	0.043***
	(0.085)	(0.067)	(0.101)	(0.003)	(0.002)
Other	0.115	0.111*	0.075	-0.027***	0.020***
	(0.083)	(0.065)	(0.098)	(0.003)	(0.002)
Tweet Category					
Corruption	0.724***	0.735***	1.144***	-0.299***	0.109***
	(0.116)	(0.092)	(0.138)	(0.004)	(0.003)
Development	2.039***	1.616***	2.506***	-0.259***	0.022***
	(0.112)	(0.089)	(0.134)	(0.004)	(0.003)
Inflation	-0.322***	-0.106	-0.169	-0.303***	0.072***
	(0.124)	(0.098)	(0.147)	(0.004)	(0.003)
Poverty	1.343***	1.033***	1.377***	-0.286***	0.104***
	(0.114)	(0.090)	(0.136)	(0.004)	(0.003)
Technology	1.593***	1.183***	1.939***	-0.252***	0.016***
	(0.112)	(0.089)	(0.134)	(0.004)	(0.003)
Friends Count	-0.286***	-	-	-	-
	(0.060)				
Followers Count	0.260***	0.336***	0.469***	0.013***	-0.007***
	(0.047)	(0.036)	(0.054)	(0.002)	(0.001)
Statuses Count	0.108	-0.009	-0.276***	-0.018***	0.006***
	(0.066)	(0.046)	(0.069)	(0.002)	(0.002)
Constant	11.193***	10.819***	13.199***	0.449***	-0.00001
	(0.282)	(0.214)	(0.323)	(0.010)	(0.007)
Observations	8,116	8,116	8,116	8,116	8,116
R^2	0.080	0.072	0.073	0.519	0.254
Adjusted R^2	0.079	0.071	0.072	0.518	0.253

4.2.2 Language Complexity. Going beyond hashtag similarity measures at high level, we also assessed the complexity of language for each category of tweets using 3 measurements of readability: Flesch-Kincaid Reading grade, Smog Index, and Gunning Fog Index [70] [7]. These metrics were used by many prior studies focused on studying the complexity of political communications [31, 57]. Here, for each original English tweet with more than 2 word tokens

---

[7]For all 3 metrics, a higher score indicates that the reading material is more complex. For instance, a Flesch-Kincaid reading grade of 10.0 suggests that the text corresponds to 10th grade reading-level.

Figure 2: Overall Compound Sentiment Score for Each Category of Tweets
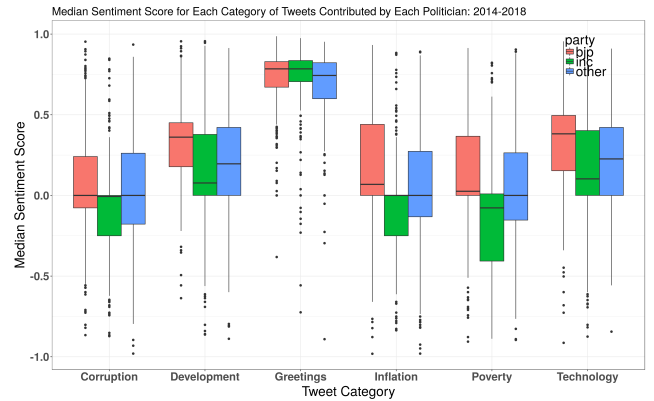
i, we use the ReadCal python library to compute all three scores and denote as the value as $readability_{i,flesch}$, $readability_{i,smog}$, $readability_{i,gunning}$. Then, we run OLS treating $readability$ as the dependent variable, and $category_i$, and $party_i$ as the independent variables.

As shown in Table 4, compared to `greetings`, all other categories of tweets except for `inflation` are correlated with higher language complexity (e.g. `development` tweets have roughly 2-grades higher reading-level than `greetings`), suggesting that tweets about substantive issues such as corruption, development, poverty and welfare are associated with higher language complexity. Further, we see that INC and 3rd party politicians contribute tweets with higher reading-levels. While a clear conclusion cannot be drawn from this alone, the suggestion is that the BJP exhibits a simpler messaging approach using more straightforward language.

4.2.3 Sentiment Appeal. Here, we first used VADER (Valence Aware Dictionary and sEntiment Reasoner), a sentiment analysis library [36] specifically attuned to social media and short texts, to generate positive, negative, and the compound affect sentiment scores of each tweet. This library is adopted by many related research [23, 50, 69] due to it has better performance compared to similar tools such as LIWC [36], as well as its ease of use. To assess VADER's performance specifically on our dataset, here we randomly sample 50 tweets with positive sentiment scores higher than negative scores, we observe 41 tweets are positive; similarly, we randomly sample 50 tweets with higher negative scores, 40 out of which are indeed negative. Looking at the visual representation of the differences between each category of tweets in Figure 2, we can see that median affect scores are more positive for both `greetings` and `technology`, indicating that politicians, as a whole, tend to utilize more positive emotional appeals when discussing celebrations or information technology. Furthermore, we also saw that INC demonstrates noticeable negativity in their tweets about corruption, inflation, and poverty comparing to BJP and 3rd party politicians. Next, for each tweet i, we apply OLS with pos_i and neg_i as the dependent variables (pos_i and neg_i are the positive, negative sentiment scores of i), and $type_i$ and $party_i$ as the independent variables.

As shown on Table 4, all substantive topics are associated with lower positive and higher negative sentiment when comparing to `greetings`. Additionally, INC and third party politicians demonstrate lower positivity and higher negativity compared to the BJP. This result aligns with prior work which suggests that politicians from non-governing parties collectively spend more effort criticizing the policies and leaders of the current majority ruling party [39]. Furthermore, previous India-specific analysis has also shown a preference for non-controversial tweeting as part of the BJPs online strategy [58].

4.3 Qualitative Analysis of the Two Key Party Leaders

Finally, we qualitatively analyzed the characteristics of tweets from key politicians: Narendra Modi and Rahul Gandhi [8]. While such tweets are a small sample, they are useful indicators of the politician's branding as well as what drives their popularity online. For both politicians, we took one representative viral tweet from each category to compare the differences.

`Development`: As demonstrated in Figure 3, while Modi talks about development in a tweet with positive valence, presenting an optimistic vision of India in a tweet that congratulates the average Indian citizen, Rahul Gandhi includes a link to a news story alongside a confrontational tweet attacking the government for a development-related fiscal policy. This contrast in styles also underlines the overall difference between the more positive-toned governing party and the more confrontational opposition party approaches.

`Technology`: In Figure 3, we show a viral tweet from Prime Minister Modi from 2016 urging young Indians to use more digital technology for financial transactions as part of the demonetization effort by the Indian government. The tone is again optimistic, aligning with the techno-optimistic style that Modi has been known for [61]. In contrast, the viral tweet from Rahul Gandhi, shown in Figure 3, includes a link to a story that claims to expose the prime minister's role in a data breach that exposed the personal information of millions of Indians. Unlike the positive tone on technology that Modi has, Gandhi presents a dystopian view of technology, of companies tracking citizens through their mobile devices. While the tweet refers to technology, the underlying intent is to attack the prime minister.

`Poverty`: On poverty, as shown in Figure 3, Modi highlights the role of his supporters in a political victory "for the poor", and has it alongside an image of himself striking a pose. The viral tweet from Rahul Gandhi uses the technical term "real wages" and adds facts to support his attack on the prime minister's economic policies. He then uses wordplay ("Modi Made Disaster" - MMD) in sarcasm.

`Inflation`: Modi's inflation tweets on average received fewer retweets than his other tweets; his most retweeted message offers discounts on petrol for cashless buyers, which is arguably more about demonetization than the price of gas. As seen in Figure 3, Modi also uses a strategy he employs frequently, of posting a picture of himself alongside the discount announcement. Rahul Gandhi,

in comparison, tweets frequently and aggressively about inflation, which has been a hot-button topic for the opposition. Here, referring to Modi as a "king of misinformation", Gandhi uses excerpts from Modi's own speech to discredit his claims on inflation.

`Corruption`: Modi tweeted significantly on corruption when he was an opposition politician. However, being in power makes that difficult since one would presumably be responsible for corruption. His highly retweeted message, in Figure 3, was a quasi-greeting to citizens. Following the Rafale arms deal scandal, corruption has been a major topic for Rahul Gandhi. He referred to Modi as "Supreme Leader" on his tweets, suggesting that the prime minister had dictatorial tendencies. The tweet highlighted here is reflective both of the frontal, insulting tone towards Modi, and of Gandhi's own appropriation of the soldier metaphor, borrowed from his opponent the long-successful value of bringing patriotism into an antagonistic exchange.

Greetings: Modi's tweets are overwhelmingly positive and usually around festivals or key moments. While Rahul Gandhi also posts positive, non-confrontational greeting tweets, he has also used congratulations on occasion with sarcasm. As shown, Modi's tweet from early 2018 congratulates the blind cricket team for its victory. His tweet uses upbeat language to present the disabled sportspersons as being a source of inspiration. In contrast, Rahul Gandhi uses combative language and irony in congratulating Narendra Modi's chief ally, the BJP party president Amit Shah, specifically calling out a financial concern that had destructive outcomes for some Indians (Figure 3).

Although these tweets are a small selection from the large body of tweets from both leaders, their virality lends them important symbolic value. The common thread is a tone of combativeness and sarcasm from Rahul Gandhi, compared to the positive-toned tweets from the Prime Minister. It is important to note, however, that all of the viral tweets came from a period when Modi has been in power. If the roles were reversed, we may well see Modi being the more aggressive one.
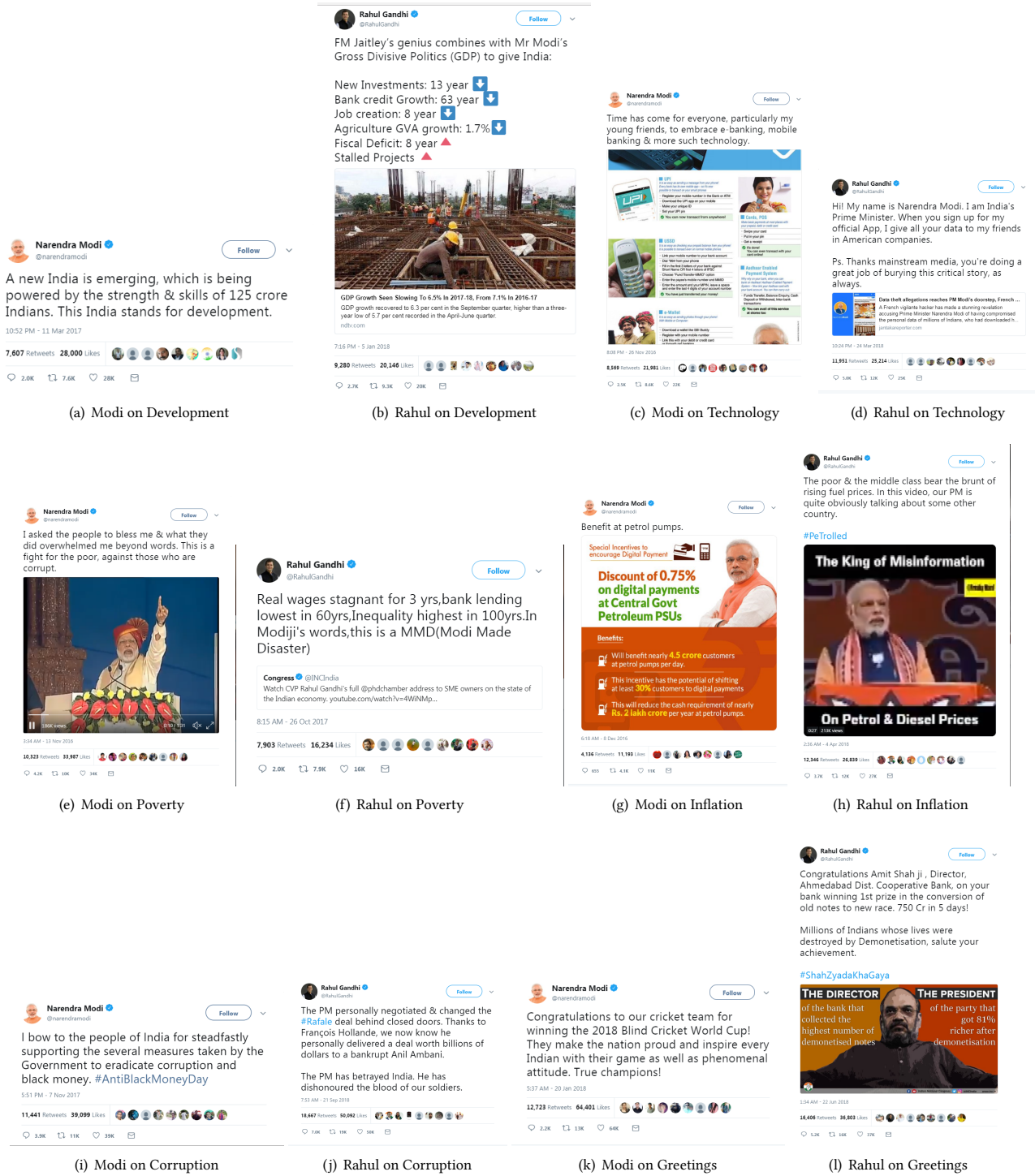
5 DISCUSSION

In this paper, we first classified a large list of Indian politicians. We then presented a picture of the social media behaviors of Indian politicians when tweeting various topics. We presented these—`technology`, `development`, `poverty`, `corruption`, and `inflation`—as policy-centric, whereas the comparative category of `greetings` is relatively less policy-centric. An important motivator of this comparison was to contextualize past research [15, 27, 29, 32, 34, 41, 55] that has suggested a dramatic shift in the media environment for politicians and political parties in terms of their choice of subjects of discussion, their use of various platform affordances (e.g., hashtags) and language strategies (e.g., sentiment), as well as their evolving interactions with each other.

Focusing on the distribution of tweet frequency across all topics, we observed that politicians do indeed more frequently tweet about greetings and celebrations as compared to policy-relevant issues. This confirmed past research findings that politicians are focused more on building soft, personal ties rather than broadcasting their

---

[8]Out of the top 200 most retweeted tweets for each category, Modi alone accounted for 41% of the total tweets and Rahul Gandhi's tweets accounted for 37%. This highlights the importance of these two key leaders in the political social media universe.

Figure 3: Narendra Modi and Rahul Gandhi's top tweets on `development, technology, poverty, inflation, corruption, and greetings`

(a) Modi on Development

(b) Rahul on Development

(c) Modi on Technology

(d) Rahul on Technology

(e) Modi on Poverty

(f) Rahul on Poverty

(g) Modi on Inflation

(h) Rahul on Inflation

(i) Modi on Corruption

(j) Rahul on Corruption

(k) Modi on Greetings

(l) Rahul on Greetings

policies. The mapping of specific politicians with repeat engagement on specific topics including corruption, poverty, and development intuitively suggests that active engagement in these areas is related to the domains of practice of these politicians. Further, politicians in power significantly lean towards the less controversial topics. For instance, the messaging of the prime minister—in particular his avoidance of the inflation topic altogether—hints at this important trend. In other words, while social media may present the spectre of political leaders constantly communicating with the general public, the lack of professional journalists engaging with political leaders affords them the ability to pick and choose topics to discuss. This may present serious challenges to the free media-led checks and balances process. Overall, we found that BJP politicians demonstrated a clear dominance in setting the agenda with many of these topical areas in part through their sheer numbers on social media.

By using text analysis, we illustrated that policy-focused tweets tend to have much higher language complexity. Additionally, by comparing the BJP and INC on corruption-related hashtags, we also found that parties in opposition are more likely to use personal or issue-based attacks on corruption, rather than talk about it in a relatively generic fashion. Moreover, we also observed higher use of nationalistic greetings hashtags in BJP's positive-themed tweets. On aggregate, we argue that politicians' preferential use of hashtags of particular sentiments in each topic clearly depicted party-based collaborative efforts in narrative-building. Next, by using qualitative methods, we saw that tweets about what would seem to be "policy" topics such as technology or inflation can sometimes a priori be examples of political performance, or confrontation, rather than substantive discussion about policy. Likewise, we also saw that greetings may not necessarily be benign or free of political implication. Finally, a deeper look at the tweets of the two most important leaders in our sample in terms of following and retweet activity—Narendra Modi and Rahul Gandhi—is instructive on both their individual styles and on the differences in sentiment between governing and opposition parties. While the governing party politician steers clear of controversy and gets widely retweeted for his non-combative, positive tweeting, the opposition politician is rewarded for being aggressive and confrontational on social media.

In sum, the engagement of politicians on casual and non-issue tweeting, and the evidence that messaging tends to be relatively straightforward, requiring less intellectual engagement from the reader, suggests that politicians gain more reward by leaning towards emotional and performative rather than substantive content. This is arguably more true if there isn't a tradition of press briefings or an empowered journalist corps asking questions to power, as is the case of India. This has ramifications for the substance of democratic communication. Leaders can appear to be communicative and responsive when they largely indulging in public relations. In particular, during times of crisis, such as right after demonetization (see Figure 4(e)), it can be deeply consequential if social media allow a leader to indulge in populist braggadocio rather respond substantively to issues. These are timely considerations during a global shock such as COVID-19, in which the demands of political actors to respond credibly to citizens' anxieties and concerns for concrete steps may be shortchanged by what emerges as the political culture of platitudes.

There are several caveats in our study worth noting here. First, the study is about India, as such some of the specifics may not apply outside of the unique context, as is often true for literature based on western cases. Nevertheless, we argue that India is the largest democracy and, given the scale of social media use in the election, this is a critical case for close examination. Second, not everyone is on Twitter. Given that India still has a substantial share of its population without access to technology or functional literacy, any study of politics and social media is likely to be biased towards those who are technology savvy and belong in higher economic classes [25]. Furthermore, other technologies such as WhatsApp and Facebook have important implications because of their widespread use among younger populations in India. However, these are largely private. The goal of our work is to study the public and performative discourse of politicians themselves, for which Twitter is a uniquely powerful tool. Equally importantly, Twitter has widespread second-order effects, in that the mainstream media as well as other social media pick up and engage with material that first appears on Twitter, which is known to be the primary form of output for many of the top political figures including Modi, Gandhi, and Shah. While we provided an in-depth assessment of politicians' Twitter priorities and strategies, our paper did not define or address the issue of success with these strategies. Future work should focus on building normative measurements of politicians' success on Twitter. One possible metric could be how many new followers, retweets or mentions politicians gain by, for instance, providing higher emotional appeal in their tweets or engaging more with prominent news media account on Twitter for different type of issues. We also did not delve deep into politicians' social media strategies within different time periods (e.g. before, during, and after elections). Future work that makes this distinction can provide valuable insights into how politicians and politically inclined individuals shift their personal branding with time. Lastly, we have invested significant effort into identifying Indian politicians at scale, and building pipelines for long-term data streaming and storage (e.g. create a database to store tweets). This dataset can be valuable to study additional ICTD-related research questions (e.g., focusing on technology, how do politicians discuss renewable energy compared to computer science?). While it would be difficult for researchers to replicate this process, we are happy to share the dataset.

6 APPENDIX

6.1 Additional Material on the Data Section

6.1.1 Twitter Account Classification: We used features including user follower count, friend count, tweet count, etc. Additionally we also used a *Bag of Words* [38] model, implemented in the *CountVectorizer* package provided by *scikit learn*, to represent the *profile description*. Furthermore, we also experimented with 5 different classification algorithms: *NaiveBayes*, *RandomForest*, *StochiasticGradientDescent*, *LogisticRegression*, and *SupportVectorMachine*. Additionally, for each classifier, we apply *GridSearchCV* to find the optimal parameters. As shown on Table 5, we find that the *SVM* classifier with an *rbf* kernel is the best. It has an accuracy score of 0.898, recall score of 0.847, precision score of 0.930, and an F1 score of 0.887.

Table 5: Performance metrics of different classifiers.

Classifier	Accuracy	Precision	Recall	F1 score
RandomForest	0.805	0.901	0.673	0.770
NaiveBayes	0.817	0.737	**0.969**	0.837
LogisticRegression	0.898	0.918	0.860	0.887
StochiasticGD	0.894	0.910	0.860	0.882
SVM (linear)	**0.907**	0.904	0.898	**0.901**
SVM (rbf)	0.898	**0.930**	0.847	0.887
SVM (poly)	0.898	0.924	0.853	0.887

6.1.2 Tweet Classification: We used *gensim*, an open source natural language processing (NLP) toolkit in python, to produce a 300-dimensional vector space representation of our corpus. Here, our input is the 4.6 million English tweets and each tweet is a single document. Further, we use *gensim*'s default parameter settings. Within this space, each unique word is assigned a numeric vector. Next, for each seed word in a category (e.g. `Technology`), we first identified the top 25 other words that had the highest cosine similarity scores to the given seed word, using *gensim*'s native function, and then manually selected the ones that made contextual sense. We repeat this process for all categories. The complete lists of keywords are in Table 6.

6.2 Additional Material on the Analyses Section

6.2.1 Bot Detection: At the time of the writing, Narendra Modi has 45.5M followers and Rahul Gandhi 8.5M. We first queried Twitter API for 1.65M and 975K most recent followers of Modi and Gandhi (we were not able to obtain the entire follower lists due to Twitter API's rate limit). Then, we randomly sampled 5K accounts from each follower list [9] and used Botometer [24] to determine their bot scores. Out of the 10K accounts, 38.4% had not posted a single tweet and 2.5% had set their accounts to private, therefore we were not able to determine if they were bots. For each of the remaining accounts, Botometer returned a score between 0 and 1 with a higher value indicting greater likelihood of being a bot. Here, an account was labeled as a bot if its score was above 0.5.

6.2.2 Identify News Media Accounts: We identify influential news media Twitter accounts using the following procedure: i) We first determined the entire list of accounts being retweeted or mentioned by politicians, and filtered out all the accounts with less than 1K followers. ii) We then generated a list of keywords that are associated with news media and careers relating to journalism such as "news", "reporter", "blogger" [10]. iii) We obtained a sublist of accounts that contain at least 1 of the keywords in the user profile description field. iv) Finally, we manually examine each account and remove the non-media accounts. A similar approach was used by Priante et al. [63] in identifying the occupations of Twitter users.

---

[9] The sample size for a population of 1.65M with 99% confidence-level and +/-3 confidence internal is 1848. Thus, a sample size of 5K is more than sufficient here.

[10] The complete list of keywords are: news, editor, journalist, columnist, magazine, blogger, anchor, newspaper, journalism, correspondent, coverage, radio, reporter, press, network, commentator.

REFERENCES

[1] L.A. Adamic and N. Glance. 2005-08. The political blogosphere and the 2004 US election: divided they blog. In *Proceedings of the 3rd international workshop on Link discovery*. 36âĂŞ43.

[2] S. Ahmed, , and K.Jaidka J. Cho. 2016. The 2014 Indian elections on Twitter: A comparison of campaign strategies of political parties. *Telematics and Informatics* 33, 4 (2016), 1071âĂŞ1087.

[3] Saifuddin Ahmed, Kokil Jaidka, and Jaeho Cho. 2016. The 2014 Indian elections on Twitter: A comparison of campaign strategies of political parties. *Telematics and Informatics* 33, 4 (2016), 1071−1087.

[4] S.and M.M.Skoric Ahmed. 2014. My name is Khan: the use of Twitter in the campaign for 2013 Pakistan General Election. In *47th Hawaii International Conference on System Sciences (HICSS), IEEE*.

[5] Eric PS Baumer, Mark Sueyoshi, and Bill Tomlinson. 2011. Bloggers and readers blogging together: Collaborative co-creation of political blogs. *Computer Supported Cooperative Work (CSCW)* 20, 1-2 (2011), 1−36.

[6] W.L. Bennett. 2012. The personalization of politics: Political identity, social media, and changing patterns of participation. *The ANNALS of the American Academy of Political and Social Science* 644, 1 (2012), 20âĂŞ39.

[7] W.L. Bennett and B. Pfetsch. 2018. Rethinking political communication in a time of disrupted public spheres. *Journal of Communication* 68, 2 (2018), 243âĂŞ253.

[8] B. Bimber. 2014. Digital media in the Obama campaigns of 2008 and 2012: Adaptation to the personalized political communication environment. *Journal of Information Technology and Politics* 11, 2 (2014), 130âĂŞ150.

[9] Troels Bøggild and Helene Helboe Pedersen. 2018. Campaigning on behalf of the party? Party constraints on candidate campaign personalisation. *European Journal of Political Research* 57, 4 (2018), 883–899.

[10] Porismita Borah. 2016. Political Facebook use: Campaign strategies used in 2008 and 2012 presidential elections. *Journal of Information Technology & Politics* 13, 4 (2016), 326–338.

[11] Ergin Bulut and Erdem Yörük. 2017. Mediatized Populisms| Digital Populism: Trolls and Political Polarization of Twitter in Turkey. *International Journal of Communication* 11 (2017), 25.

[12] Z.M. Bureau. 2018. PM Narendra Modi tells BJP MPs to have at least 3 lakh followers on Twitter.

[13] Jennifer Bussell. 2012. *Corruption and reform in India: Public services in the digital age*. Cambridge University Press.

[14] Rosie Campbell, Philip Cowley, Nick Vivyan, and Markus Wagner. 2016. Legislator dissent as a valence signal. *British Journal of Political Science* (2016), 1−24.

[15] A. Casero-RipollÃŠs, , and M.Sintes-Olivella P. Franch. 2017. "The populist political communication style in action: Podemos's issues and functions on Twitter during the. *American Behavioral Scientist* 61, 9 (2017), 986âĂŞ1001.

[16] A. Chadwick. 2017. *The hybrid media system: Politics and power*. Oxford University Press.

[17] Andrew Chadwick and Jennifer Stromer-Galley. 2016. Digital media, power, and democracy in parties and election campaigns: Party decline or party renewal?

[18] Kanchan Chandra. 2016. *Democratic dynasties: State, party, and family in contemporary Indian politics*. Cambridge University Press.

[19] M. Conover, J. Ratkiewicz, M.R. Francisco, B. GonÃğalves, F. Menczer, and A. Flammini. 2011. Political polarization on twitter. *Icwsm* 133 (2011), 89âĂŞ96.

[20] J.C.and J.E.Camargo Correa. 2017. "Ideological consumerism in colombian elections. *Cyberpsychology, Behavior, and Social Networking* (2017).

[21] N. Curato. 2017. Flirting with authoritarian fantasies? Rodrigo Duterte and the new terms of Philippine populism. *Journal of Contemporary Asia* 47, 1 (2017), 142âĂŞ153.

[22] Y.and C.Baden David. 2018. Reframing community boundaries: the erosive power of new media spaces in authoritarian societies. *Information, Communication and Society* (2018), 1âĂŞ18.

[23] Thomas Davidson, Dana Warmsley, Michael Macy, and Ingmar Weber. 2017. Automated hate speech detection and the problem of offensive language. In *Eleventh international aaai conference on web and social media*.

[24] Clayton Allen Davis, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. Botornot: A system to evaluate social bots. In *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, 273–274.

[25] Antonella De Angeli, Uday Athavankar, Anirudha Joshi, Lynne Coventry, and Graham I Johnson. 2004. Introducing ATMs in India: a contextual inquiry. *Interacting with Computers* 16, 1 (2004), 29–44.

[26] Evelien DâĂŹheer and Pieter Verdegem. 2014. Conversations about the elections on Twitter: Towards a structural understanding of TwitterâĂŹs relation with the political and the media field. *European Journal of Communication* 29, 6 (2014), 720−734.

[27] G. Enli. 2017. "Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election. *European Journal of Communication* 32, 1 (2017), 50âĂŞ61.

[28] G.S.and E.SkogerbÃÿ Enli. 2013. Personalized campaigns in party-centred politics: Twitter and Facebook as arenas for political communication. *Information, Communication and Society* 16, 5 (2013), 757âĂŞ774.

Table 6: List of Keywords and Terms Used to Generate Each Category of Tweets. Note: all Hindi Tokens were Translated.

Category	Keywords
Technology	comput, internet, kiosk, scienc, scient, technolog, telecom, satellite, digital, engineer, innovati, biotech, start up, software, startup, wifi , fintech, pslv , computer, device, multimodal, biometric, msme , electrical, 4g , aviation, electronic, cisf , app , battery, isro , automobile, portal, bhimapp, digitalindia, engineering, hotspot, wi-fi, web , multi-modal, innovate, iot , technological, real-time, automatic, dashboard, automated, csir , brahmos, software, digitisation, nano , tracking, rocket, smartphones, invention, test-firing, drdo , digital, smartphone, gsat-19, bsnl , sme , start-up, world-class, electronics, msmes, biotech, gps , mfg , digitalpayments, cashless, internet, jio , railwire, digitalpayment, petrochemical, e-commerce, robotics, technology, apps , virtual, technologist, mobileapp, entrepreneur, wifi , hardware, drone, prototype, resourcesat-2a, ip , broadband, ict , scientist, orbit, makeinindia, entrepreneurship, eco friendly, tech , e-wallets, android, technical, innovation, upi , state-of-the-art, evinindia, cng , startup, google, multi-purpose, pioneer, umang, cipet, srcitizen science, nobelprize, cartosat-2, mangalyaan, scientific, petroleum, startupindia, hydrocarbon, gsat , digita, tablet, digiti, ecommerce, techno, android, aeps , choupal, pslvc38, cartosat-2, dispensaryautomotive, kampyootar, vigyaan, paaribhaashikee, vaigyaanik, kampyootar, shilp-kala, injeeniyar, praudyogikee, phesabuk, tvitar, intaranet, instraagraam, vaigyaanikon, aadhunik, ilektroniks, fesabuk, tvitar, atyaadhunik, kampyootar, vaigyaanik, ilektronik, takaneekee, intaranet, brodabaind, vigyaanan, vaeephaee, sophtaveyar, teknolojee, vebasait, sikyoritee, laeev , tveetar, injeeniyar, ee-mitr, vigyaan, klastar, phesabuk, post, kampyootar, googal
Poverty	accessib, poor , alleviation, malnutrition, financialinclusion, low-income, homeless, unemploy, handicap, amenity, prostitution, marginalized, neo-middle, mortality, worldfooddaytrapped, destitute, rightsnotcharity, casteism, neglected, exploited, inequality, backwardness, starving, illiteracy, deprivation, foodwater, robbed, upliftment, struggling, downtrodd, shortage, benefitting, govtthatcares, child labourvulnerable, untouchab, deprived, under-privileged, poorest, underserved, emancipation, casteless, downtrodden, poverty, marginali, unprivileged, exploiting, underprivileged, jobless, nrega, slumdisadvantaged, childlabour, labourer, hard-working, morbidity, survival, globalhungerindex, assisted-living, elderly, unskilled, exploitation, affordab, mathrupoorna, dwindling, remotest, hardshiplivelihood, divyang, welfare, needy, universalhealthcoverage, rural, garib, houseless, daridr, niraashrit, besahaara, nirdhan, bekas , mufali, abhaav , gareeb , nirbal, muhataaj, zarooratamand, bhookh , peedit, shoshit, kisaan, krsh , laachaar, bebas , peedit, laachaar
Development	self-reliance, iip , developmental, economy, cleanenergy, socio-economic, indiameansbusiness, gdpgrowth, gva , patenting, employability, growth, consumption, growth, urbanisation, infrastructure, year-on-year, budgetforbetterindia, double-digit, agricultural, macro-economic, globalization, trajectory, fiscal, q3 , deflation, growt, industrialization, q2 , modernity, competitiveness, expanding, infrastructural, dvpt , urbanization, macroeconomic, newindiatakeoff, accelerating, development, liquidity, budget2017, wpi , infra, devpt, psbs , investment, modernising, expenditure, productivity, mfg , accessibility, disparity, gdp , infrastruct, connectivity, dvlp , devt , devl , productivity, globalisation, protectionism, fy17 , q4 , production, upliftment, economicsurvey2018, shrinking, efficiency, crisil, inflow, slowdown, logistics, grwth, connectivity, q1 , macro, oilseed, transport, transportation, output, yoy , diversification, equality, steep, borrowing, deficit, profitability, maize, sectoral, indianeconomy, stimulate, export, krishikalyanabhiyan, ikaas, vrddhi, sanvrddhi, utpaadan, unnati, samrddhi, baagavaanee, badhotaree, petrol-deejal, khetihar, phal-sabjiyon, vrddhi, vikaas, mahangaee, badhottaree, mahangaee, pashupaalakon, utthaan, svarozagaar, utpaadakata, kushalata, keematon, vikaas., karz , pratishat, utpaadan, vrddhi., sinchaee, arthavyavastha, utpaadan, badhottaree, khetee , badhottaree, paidaavaar, unnayan, vrddhi, rojagaar, udyaanikee
Greetings	harvest, vijayadashami, sankranti, mubarakmay, b'day, republic day, wonderful, gurupurnima, utkaladibasa, christmas, cheiraoba, celebrated, tranquility, ramadan, gratitude, eiduladha, new year, warmth, celebrating, eidmubarak, festivity, xmas , mahashivaratri, makarsankranti, birthday, long life, colourful, compassion, festival, childrensday, holi , prosper, id-e-milad, bday , loving, greetngs, fantastic, peacefulness, joyful, prosperi, wishing, wish , brotherhood, birthday, happybirthday, fabulous, greeting, joy , baisakhi, bless, heartiest, gudipadwa, peace, oneness, merry, beloved, bathukamma, longlife, eidaladha, akshayatritiya, happyholi, eidulfitr, bohagbihu, joyous, harmonious, blessed, heartfull, boishakh, congradulations, congradul, congratulation, congratulat, happydiwali, delightful, reopens, puthandu, holiday, dhanteras, onam , easter, greeted, happiness, auspicious, janmashtami, sankranthi, bliss, mahavirjayanti, greets, abundance, samvatsari, colorful, praying, carnival, heartwarming, harmony, heartfelt, prosperity, cherish, contentment, blissful, pongal, terrific, jesus, divine, diwali, gudhi padwa, good health, independence day, sincerest, harvest, vijayadashami, sankranti, mubarakmay, b'day, republic day, wonderful, gurupurnima, utkaladibasa, christmas, cheiraoba, celebrated, tranquility, ramadan, gratitude, eiduladha, new year, warmth, celebrating, eidmubarak, festivity, xmas , mahashivaratri, makarsankranti, birthday, long life, colourful, compassion, festival, childrensday, holi , prosper, id-e-milad, bday , loving, greetngs, fantastic, peacefulness, joyful, prosperi, wishing, wish , brotherhood, birthday, happybirthday, fabulous, greeting, joy , baisakhi, bless, heartiest, gudipadwa, peace, oneness, merry, beloved, bathukamma, longlife, eidaladha, akshayatritiya, happyholi, eidulfitr, bohagbihu, joyous, harmonious, blessed, heartfull, boishakh, congradulations, happydiwali, delightful, reopens, puthandu, holiday, dhanteras, onam , easter, greeted, happiness, auspicious, janmashtami, sankranthi, bliss, mahavirjayanti, greets, abundance, samvatsari, colorful, praying, carnival, heartwarming, harmony, heartfelt, prosperity, cherish, contentment, blissful, pongal, terrific, jesus, divine, diwali, gudhi padwa, good health, independence day, haardik, shubh , janmadin, chhuttee ka din, avasar par, shubhakaamanaen, holee , naya saal, tyohaar, gaandhee jayantee, mahaaveer jayantee, buddh poornima, lohadee, janmadivas, raamanavamee, janmadin, baisaakhee, janamadin, pongal, punyatithi, shubhakaamanaayen, holika, shubhakaamanae, shubhakaamanaen, deepaavalee,jeevanaparv, navaraatr, shubhakaamanaen, navaraatree, shubhakaamanaen., shubh , jayantee, shubhakaamanaayen, shubhakaamanaaye, janmaashtamee, krisamas, dhanateras, shubhakaamanaen, mubaarakabaad, mahaaparv, makar , vijayadashamee, badhaiyaan, jayantee, janm , eed , shubhakaamana, janmotsav, mahaashivaraatri, janmadeen, tyauhaar, utsav, parv , makarasankraanti, shubhakaamana, shubhakaam, badhaiyaan, paavan , vijayaadashamee, rakshaabandhan, phitar, janyatee, mubaarak, mangalamay
Corruption	rafale, corruption, deal , scam , rafalescam, rafaledeal, cag , mumbaiagainstrafalescam, middleman, rafalescamsilence, corrupt, anothercongressdefencescam, congressdefencescam, hal , crony, defaulters, deals, doublestandards, scandal, cheating, transparency, quattrocchi, truthofrafale, bofors, 130000crorerafalescam, scams, dubious, exposed, iafbacksrafale, rafalerobbery, modiscandals, blackmoneycrackdown, siphoning, rafalescammonday30th, hawala, defaulter, proberafalescam, modiscam, vyapam, modiambanirafaleblockbuster, thegreatrafalemystery, kickback, blackmoney, augusta, westland, bribe, chowkidarnahibhagidar, fraudster, rafalechargesheet, jpc , embezzlement, corruption-free, fraudulent, rahulgandhiwithhal, bribery, rafalepejhoothbandkaro, agusta, agustawestland, mehulchoksi, modiambanirafalescam, rafaledeal, favouritism, aircel-maxis, pnbfraudcase, rafalescamexpose, pnbscam, modighotala, choksi, pnbfraud, dealmeinkuchkalahai, niravmodilootsindia, niravmodi, robert vadra, mehul choksi, adarsh scam, national herald, rahulkapurakhandanchor, 2g spectrum, antigua, vijay mallya, puricongresschorhai, raaphel, raphel , raphaal , raafel, bhrasht, bhrashtaachaar, jhoot , sauda , saude , agasta, vestalaind, vestalend, vestalaind, dhokha , dhoka , dhokhaadhadee, dhokaadhadee, dhokebaaz, dhokhebaaz, vyaapam, vyaapaman, vyaapam, khaoonga, ophaset, ghotaalebaajon, ghotaalon, ghotaalo, ghotaale, ghotaala, ghotaalebaaj, dalaalee, dalaal , kameeshanakhoree, kameeshan, kamishan, kameeshanakhor, kamishanakhor, kamishanakhoree, kameeshanabaazee, lokapaal, janalokapaal, poonjeepati, kaaledhan, kaalaadhan, aadarshaghotaala_sangharsh_yaatra, rishvat, rishavat, rishvatakhor, rishvatakhoree, rishvatabaazee, jhaansa, karapshan, robart, vaadra, vaadra, aadarsh ghotaala, 2g , bophors, bofors, vijay maalya, maalaya, aanteegva, enteegua, entigva, aanteegua, ghotalapradeshmp, puricongresschorhai, ghotala, rafale, vyapam
Inflation	petrol, diesel, bharatbandh, pricing, loot , fuelloot, lpg , inflation, cylinder, highdiesel, fuelhikeprice, bharatbandhinflation, petrolpricehike, mehngaikiaag, rocketed, fuelonfire, crude, fuellootbysuitboot, wpi , petrolchormodi, inflation, petroldieselpricehike, daam , silendar, petrol, deezal, petrol-deezal, mahangaee, silendar, silindar, daamon, petroladeejal, paitrol, rasoee , elapeejee, manhagaee, mahangaee

[29] Heather K Evans, Victoria Cordova, and Savannah Sipole. 2014. Twitter style: An analysis of how house candidates used Twitter in their 2012 campaigns. *PS: Political Science & Politics* 47, 2 (2014), 454–462.

[30] Tom Feltwell, John Vines, Karen Salt, Mark Blythe, Ben Kirman, Julie Barnett, Phillip Brooker, and Shaun Lawson. 2017. Counter-Discourse Activism on Social Media: The Case of Challenging âĂIJPoverty PornâĂİ Television. *Computer Supported Cooperative Work (CSCW)* 26, 3 (2017), 345–385.

[31] I. Flaounas, O. Ali, T. Lansdall-Welfare, T. De Bie, N. Mosdell, J. Lewis, and N. Cristianini. 2013. Research methods in the age of digital journalism: Massive-scale automated analysis of news-contentâĂŤtopics, style and gender. *Digital Journalism* 1, 1 (2013), 102âĂŞ116.

[32] G. Fuller, , and A.Jolly C. Fisher. 2018. Malcolm Turnbull's conversational career on Twitter: the case of the Australian Prime Minister and the NBN. *Media International Australia* 167, 1 (2018), 88âĂŞ104.

[33] D. Gayo-Avello. 2017. Social Media Won't Free Us. *IEEE Internet Computing* 21, 4 (2017), 98âĂŞ101.

[34] Libby Hemphill, Jahna Otterbacher, and Matthew Shapiro. 2013. What's congress doing on twitter?. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 877–886.

[35] P.N. Howard. 2006. *New media campaigns and the managed citizen.* Cambridge University Press.

[36] C.J. Hutto and E.E. Gilbert. 2014-06. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In *Eighth International Conference on Weblogs and Social Media (ICWSM-14).* Ann Arbor, MI.

[37] C. Jaffrelot. 2015. Narendra Modi and the power of television in Gujarat. *Television and New Media* (2015).

[38] Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2016. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759* (2016).

[39] Andreas Jungherr. 2016. Twitter use in election campaigns: A systematic literature review. *Journal of information technology & politics* 13, 1 (2016), 72–91.

[40] Ravinder Kaur. 2015. Good times, brought to you by brand Modi. *Television & New Media* 16, 4 (2015), 323–330.

[41] D. Kellner. 2018. Donald Trump and the War on the Media: From Election '16 into the Trump Presidency. In *The Trump Presidency, Journalism, and Democracy, Routledge.* 19âĂŞ38.

[42] Yubo Kou, Yong Ming Kow, Xinning Gui, and Waikuen Cheng. 2017. One Social Movement, Two Social Media Sites: A Comparative Study of Public Discourses. *Computer Supported Cooperative Work (CSCW)* 26, 4-6 (2017), 807–836.

[43] D. Kreiss. 2016. Seizing the moment: The presidential campaigns' use of Twitter during the 2012 electoral cycle. *New Media and Society* 18, 8 (2016), 1473âĂŞ1490.

[44] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web.* AcM, 591–600.

[45] C. Lee. 2017. Facebooking to Power: The Social Media Presence of Malaysian Politicians. *ISEAS Perspectives* 74 (2017).

[46] G. Lotan, E. Graeff, M. Ananny, , and D.Gaffney I. Pearce. 2011. "The Arab Spring| the revolutions were tweeted: Information flows during the. *International journal of communication* 5 (2011), 31.

[47] M Lufkens. 2016. Twiplomacy Study 2016. *Burson-Marsteller, Editor, Geneva* (2016).

[48] Swati Maheshwari and Colin Sparks. 2018. Political elites and journalistic practices in India: A case of institutionalized heteronomy. *Journalism* (2018), 1464884918761630.

[49] S.and C.Sparks Maheshwari. 2018. Political elites and journalistic practices in India: A case of institutionalized heteronomy. *Journalism* (2018).

[50] Lydia Manikonda, Ghazaleh Beigi, Huan Liu, and Subbarao Kambhampati. 2018. Twitter for sparking a movement, reddit for sharing the moment:# metoo through the lens of social media. *arXiv preprint arXiv:1803.08022* (2018).

[51] Christopher M Mascaro and S Goggins. 2010. Collaborative information seeking in an online political group environment. In *International workshop on CIS of CSCW conference*. Citeseer.

[52] S.C. McGregor. 2018. Personalization, social media, and voting: Effects of candidate self-personalization on vote intention. *new media and society* 20, 3 (2018), 1139âĂŞ1160.

[53] L. Meeks. 2016. Gendered styles, gendered differences: Candidates' use of personalization and interactivity on Twitter. *Journal of Information Technology and Politics* 13, 4 (2016), 295âĂŞ310.

[54] E. Morozov. 2009. Iran: Downside to the "twitter revolution". *Dissent* 56, 4 (2009), 10âĂŞ14.

[55] V. Mudgal. 2015. Framing the 2014 elections: The curious absence of development. *Television and New Media* 16, 4 (2015), 354âĂŞ360.

[56] B. Ohm. 2015. Organizing Popular Discourse with and against the Media: Notes on the Making of Narendra Modi and Recep Tayyip Erdogan as Leaders-without-Alternative. *Television and New Media* 16, 4 (2015), 370âĂŞ377.

[57] B.L. Ott. 2017. The age of Twitter: Donald J. Trump and the politics of debasement. Critical Studies in Media. *Communication* 34, 1 (2017), 59âĂŞ68.

[58] Joyojeet Pal. 2015. Banalities turned viral: Narendra Modi and the political tweet. *Television and New Media* 16, 4 (2015), 378–387.

[59] J. Pal, P. Chandra, and V.V. VydiswaranâĂŹ. 2016. Twitter and the rebranding of Narendra Modi. *Economic and Political Weekly* 51, 8 (2016), 52âĂŞ60.

[60] Joyojeet Pal, Udit Thawani, Elmer van der Vlugt, Wim Out, Priyank Chandra, et al. 2018. Speaking their Mind: Populist Style and Antagonistic Messaging in the Tweets of Donald Trump, Narendra Modi, Nigel Farage, and Geert Wilders. *Computer Supported Cooperative Work (CSCW)* (2018), 1–34.

[61] J.and A.Gonawela Pal. 2016. *Political social media in the global South. Conference on e-Business, e-Services and e-Society*. Springer.

[62] Thomas Poguntke and Paul Webb. 2007. *The presidentialization of politics: A comparative study of modern democracies*. Oxford University Press on Demand.

[63] A. Priante, D. Hiemstra, T. van den Broek, A. Saeed, M. Ehrenhard, and A. Need. 2016. #WhoAmI in 160 characters? Classifying social identities based on twitter

[64] Arvind Rajagopal. 2017. On Media and Politics in India: An Interview with Paranjoy Guha Thakurta. *South Asia: Journal of South Asian Studies* 40, 1 (2017), 175–190.

[65] U. Rao. 2010. *News as culture: Journalistic practices and the remaking of Indian leadership traditions*. Berghahn Books.

[66] U.M. Rodrigues. 2014. Social media's impact on journalism: a study of media's coverage of anti-corruption protests in India. *Global media journal: Australian edition* 8, 1 (2014), 1âĂŞ10.

[67] Mirko Tobias Schäfer. 2016. Challenging Citizenship: Social Media and Big Data. *Computer Supported Cooperative Work (CSCW)* 25, 2-3 (2016), 111–113.

[68] P. Sharma and T.S. Moh. 2016-12. Prediction of indian election using sentiment analysis on hindi twitter. In *2016 IEEE International Conference on Big Data (Big Data)*. IEEE, 1966âĂŞ1971.

[69] Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee, and Huan Liu. 2018. Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media. *arXiv preprint arXiv:1809.01286* (2018).

[70] Luo Si and Jamie Callan. 2001. A statistical model for scientific readability. In *Proceedings of the tenth international conference on Information and knowledge management*. ACM, 574–576.

[71] S. Soroka, L. Young, and M. Balmas. 2015. Bad news or mad news? Sentiment scoring of negativity, fear, and anger in news content. *The ANNALS of the American Academy of Political and Social Science* 659, 1 (2015), 108âĂŞ121.

[72] Stefan Stieglitz and Linh Dang-Xuan. 2013. Social media and political communication: a social media analytics framework. *Social Network Analysis and Mining* 3, 4 (2013), 1277–1291.

[73] R. Sundaram. 2015. Publicity, transparency, and the circulation engine: the media sting in India. *Current Anthropology* 56, 12 (2015), 297âĂŞ305.

[74] F. Tusa. 2013. "How social media can shape a protest movement: The cases of Egypt in. *Arab Media and Society* 17 (2013), 1âĂŞ19.

[75] Nikki Usher, Jesse Holcomb, and Justin Littman. 2018. Twitter makes it worse: Political journalists, gendered echo chambers, and the amplification of gender bias. *The International Journal of Press/Politics* 23, 3 (2018), 324–344.

[76] Onur Varol, Emilio Ferrara, Clayton A Davis, Filippo Menczer, and Alessandro Flammini. 2017. Online human-bot interactions: Detection, estimation, and characterization. *arXiv preprint arXiv:1703.03107* (2017).

[77] J. Wilkerson and A. Casas. 2017. Large-scale computerized text analysis in political science: Opportunities and challenges. *Annual Review of Political Science* 20 (2017), 529âĂŞ544.

[78] J.L. Zittrain, R. Faris, H. Noman, J. Clark, , and C.Tilton R. Morrison-Westphal. 2017. The Shifting Landscape of Global Internet Censorship.

profile descriptions. In *Proceedings of the First Workshop on NLP and Computational Social Science*. 55âĂŞ65.